

BUNDESREPUBLIK DEUTSCHLAND



REC'D 04 JAN 2005

WIPO PCT

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 103 50 474.5

Anmeldetag: 23. Oktober 2003

Anmelder/Inhaber: Universität Leipzig, 04109 Leipzig/DE

Bezeichnung: Verfahren zur Selektion von Biomolekülen aus
Varianten-Bibliotheken von Biomolekülen

IPC: C 12 Q 1/68

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 22. November 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

[Handwritten signature]
Brosig

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

3



Verfahren zur Selektion von Biomolekülen aus Varianten-Bibliotheken von Biomolekülen

Die Erfindung betrifft ein Verfahren zur Selektion von Biomolekülen aus Varianten-Bibliotheken von Biomolekülen, insbesondere von Enzymen oder anderen biokatalytisch aktiven Biomolekülen. Biomoleküle finden vielseitige Verwendung in technologischen oder medizinischen Anwendungen und Prozessen. Viele der dafür benötigten Eigenschaften der Biomoleküle sind in der Natur so nicht vorhanden oder noch nicht identifiziert worden. Die Generierung solcher neuen Eigenschaften aus vorhandenen Biomolekülen erfordert die Herstellung sehr großer Varianten-Bibliotheken mit zufällig veränderten Zusammensetzungen. Die Identifizierung der Varianten mit den gesuchten Eigenschaften erfordert geeignete Selektions- oder Durchmusterungs-Verfahren.

Neue Biomoleküle können durch die Verknüpfung der neuen Eigenschaft an das Überleben oder einen hinreichend großen Wachstumsvorteil eines Organismus erzeugt werden. Hierbei wird die Varianten-Bibliothek in einen entsprechenden Organismus überführt und die Wachstumsbedingungen so gewählt, dass nur die Individuen überleben oder vergleichsweise schneller wachsen, welche eine Variante des Biomoleküls mit der gesuchten neuen Eigenschaft produzieren (Zaccolo M, Gherardi E. The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. J. Mol. Biol. 1999. 285, 775-83. oder Samuelson JC, Xu SY. Directed evolution of restriction endonuclease BstYI to achieve increased substrate specificity. J. Mol. Biol. 2002. 319,673-83). Diese Anwendung ist nur auf die Selektion eines eng begrenzten Kreis von Biomolekülen anwendbar, die einem gewählten Organismus einen Vorteil verschaffen können. Biomoleküle, welche beliebige chemische Reaktionen katalysieren, sind so nicht selektierbar. Da der Organismus über den gesamten Selektionszeitraum am Leben bleiben muss, sind toxische oder anderweitig für ein Wachstum nachteilige Eigenschaften nicht selektierbar.

Eine weitere Methode zur Selektion neuer Biomoleküle ist die Verknüpfung des Biomoleküls mit der codierenden Nukleinsäure-Sequenz (Amstutz P, Forrer P, Zahnd

- C, Pluckthun A. In vitro display technologies: novel developments and applications. Curr. Opin. Biotechnol. 2001. 12. 400-5. Xia G, Chen L, Sera T, Fa M, Schultz PG, Romesberg FE. Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. Proc. Natl. Acad. Sci. USA. 2002. 99. 6597-602. Pschorr J. Genotyp und Phänotyp koppelnde Verbindung. DE0019646372C1). Eine Anwendung dieser Technologien mit lebenden Organismen wie Phagen oder Bakterien beschränkt das Spektrum wiederum auf nicht toxische oder nicht wachstumshemmende Biomoleküle. Ebenso dürfen Substrate und Produkte der gesuchten Reaktion auf den präsentierenden Organismus nicht schädigend einwirken.
- 5 10 Zusätzlich lassen sich katalytische Aktivitäten nur selektieren, wenn sich Biomolekül und Substrat an dem gleichen Organismus präsentieren lassen. Da sich die Aktivität der katalytischen Biomoleküle nicht auf den sie präsentierenden Organismus begrenzen lässt und diese somit auch Reaktionen an anderen Individuen der Bibliothek stattfinden können, führt diese Methode häufig zur Falsch-Selektion von Biomolekülen.
- 15 Bei den Durchmusterungs-Verfahren (Screening-Verfahren) wird jede Variante einer Biomolekül-Bibliothek einzeln hinsichtlich der gesuchten Eigenschaft untersucht. (Joo H, Lin Z, Arnold FH. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. Nature. 1999. 399. 670-3. Korbel GA, Lalic G, Shair MD. Reaction
- 20 microarrays: a method for rapidly determining the enantiomeric excess of thousands of samples. J. Am. Chem. Soc. 2001. 123. 361-2.) Selbst bei sehr kurzen Messzeiten (z. B. 100 ms pro Variante) erfordert diese Methode einen sehr hohen Zeitaufwand (z.B. 22 Tage) für die Untersuchung großer Bibliotheken (z.B. 10^7). Das kontinuierliche Messen von Varianten in diesen Größenordnungen erfordert den Aufbau entsprechend
- 25 komplexer Apparate. Außerdem muss für jede Variante der Bibliothek ein entsprechender Eigenschaftstest durchgeführt werden, was zu sehr hohen Kosten dieser Methoden führt.

Aufgabe der vorliegenden Erfindung ist es ein Verfahren zur Identifizierung von

30 Biomolekülen in Varianten-Bibliotheken von Biomolekülen anzugeben.

Erfindungsgemäß wird die Aufgabe gelöst durch ein Verfahren zur Selektion von Biomolekülen aus Varianten-Bibliotheken von Biomolekülen umfassend die Schritte:

- a.) Herstellung einer Varianten-Bibliothek, bestehend aus einer Anzahl (B_0) von Varianten, der für das Biomolekül codierenden Gensequenz, und
- 5 b.) Aufteilung der Variantenbibliothek in eine Anzahl (W_0) von Kompartimenten, die mindestens um einen Faktor 10, bevorzugt einen Faktor 100, kleiner ist, als die Anzahl der in der Varianten-Bibliothek enthaltenen Varianten,
wobei jedes Kompartiment eine Teilbibliothek enthält, die $K_0=B_0/W_0$ Varianten enthält,
- 10 c.) Produktion von Biomolekülen in den Kompartimenten, und Test der in den einzelnen Kompartimenten erhaltenen Biomoleküle auf eine bestimmte Eigenschaft, bevorzugt eine biokatalytische Aktivität,
- d.) Auswahl mindestens eines Kompartiments, in dem Biomoleküle enthalten sind, welche die gewünschten Eigenschaften erfüllen,
- 15 e.) Aufteilung der in dem ausgewählten Kompartiment enthaltenen Teilbibliothek in weitere Kompartimente und n-faches Wiederholen der Schritte c.) bis e.) bis in jedem Kompartiment nur noch maximal eine Variante ($K_n \leq 1$) der für das Biomolekül codierenden Gensequenz enthalten ist.
- 20

Dieses Verfahren ist insbesondere zur Generierung von Biomolekülen mit neuen katalytischen Aktivitäten geeignet, die entweder in der Natur gar nicht vorkommen oder zumindest von dem gewählten Ausgangs-Biomolekül so nicht katalysiert werden.

25 Außerdem können mit diesem Verfahren vorhandene katalytische Aktivitäten an äußere Bedingungen, wie z.B. Temperatur oder Lösungsmittel, angepasst werden, unter denen bisher keine oder nur verschwindend geringe Aktivität vorhanden war.

Da eine Produktion der Biomoleküle in der vorliegenden Erfindung zu einem Absterben der Organismen führen kann oder durch zellfreie Systeme erfolgen kann, kann das

30 Verfahren auf alle Arten von Biomoleküle angewandt werden und ist nicht auf nicht

- 5 toxische oder nicht wachstumshemmende Aktivitäten begrenzt. Da bis zu einer Million oder mehr Varianten mit einem Test und gleichzeitig auf die entsprechende Eigenschaft hin untersucht werden, reduziert sich die für das Durchmustern der Bibliothek benötigten Zeit und die für die Eigenschaftstests benötigten Kosten um einen entsprechenden Faktor. Varianten, welche die gesuchten Eigenschaften aufweisen, können sicher und reproduzierbar aus den ursprünglichen Varianten-Gemischen isoliert werden.

Im Schritt a.) des Verfahren wird eine Varianten-Bibliothek der für das Biomolekül codierenden Gensequenz mittels Standardmethoden der Molekularbiologie hergestellt.

10

Unter Variantenbibliothek wird im Sinne des vorliegenden Erfindung verstanden: Mischung aus Proteinen oder Nukleinsäuren, welche sich in mindestens einer Position in ihrer Sequenz voneinander unterscheiden.

15

Bevorzugt besteht die Variantenbibliothek bestehend aus einer Anzahl von Varianten in einem Größenbereich von $B_0 = 10^3$ bis $B_0 = 10^{15}$. So können zum Beispiel in einen Teilbereich des Biomoleküls zufällig gewählte Sequenzbausteine eingeführt werden, so dass im Fall einer Nukleinsäure bei 25 geänderten Positionen eine Bibliotheksgröße von $4^{25} = 1,1 \times 10^{15}$ oder im Fall eines Proteins bei 7 geänderten Positionen eine Bibliotheksgröße von $20^7 = 1,3 \times 10^9$ entstehen kann.

20

Besonders bevorzugt liegt der Größenbereich zwischen $B_0 = 10^5$ bis $B_0 = 10^9$.

Besonders bevorzugt besteht die Variantenbibliothek aus DNA-Plasmiden oder linearen Nukleinsäuremolekülen, welche die für das Biomolekül codierende Gensequenz enthalten.

25

Biomoleküle im Sinne des vorliegenden Erfindung sind Proteine, Nukleinsäuren oder andere aus organischen Bausteinen bestehende Biopolymere. Bevorzugt sind diese Biomoleküle, Enzyme oder Ribozyme oder andere Biomoleküle, die als Biokatalysatoren, die Umsetzung chemischer oder biochemischer Stoffe beschleunigen können.

Standardmethoden der Molekularbiologie mit denen eine solche Variantenbibliothek hergestellt werden kann, sind beispielsweise fehlerhafte Vervielfältigungsmethoden für Nukleinsäuren. Hierfür werden replizierende Enzyme, z.B. Polymerasen, welche die Neusynthese eines Biomoleküls mit Hilfe einer Matrize durchführen, angewandt. Die Einführung von Fehlern und somit die Erzeugung von unterschiedlichen Varianten erfolgt entweder durch die natürlich vorhandene Fehlerrate dieser replizierenden Enzyme oder kann durch Veränderung der Reaktionsbedingungen (z.B. Ungleichgewicht der Synthesebausteine, Zugabe von Baustein-Analoga, Veränderung der Puffer-Zusammensetzung) erhöht werden. Neben der Einführung von Fehlern kann eine Variantenbibliothek unter Ausnutzung der natürlich vorkommenden Diversität für ein bestimmtes Biomolekül oder eine Biomolekülklasse erzeugt werden.

Im Schritt b.) wird die Variantenbibliothek in eine Anzahl W_0 von Kompartimenten, die mindestens um einen Faktor 10, bevorzugt einen Faktor 100, kleiner ist, als die Anzahl der in der Varianten-Bibliothek enthaltenen Varianten, aufgeteilt.

Hierbei kann vor der Aufteilung die Überführung der Variantenbibliothek in einen Organismen erfolgen oder die Aufteilung auf Ebene der codierenden Sequenzen erfolgen. Die Aufteilung erfolgt so, dass jede Variante der Bibliothek mindestens einmal, bevorzugt genau einmal, vorkommt.

Die Produktion (Expression) der Biomoleküle kann durch den Organismus oder durch *in vitro* Expressionssysteme (z.B. Zellextrakte) erfolgen.

Als Expressions-Organismus kommen alle in der Molekularbiologie standardmäßig zu Expression von Biomolekülen, wie Proteinen, verwendeten Organismen in Betracht, der Expressionsorganismus wird dabei in Anhängigkeit des zu exprimierenden Biomoleküls gewählt. Bevorzugte Expressionsorganismen sind bakterielle Zellen (z.B. *E. coli*, *B. subtilis*) oder eukaryontische Zellen (z.B. *S. cerevisiae*, Insektenzellen, Tumorzellen).

Durch die Überführung der Variantenbibliothek in den Expressionsorganismus entstehen einzelne Klone des Expressionsorganismus. Dabei enthält ein Klon jeweils einen definierten Genotyp, d. h. eine Variante der für das Biomolekül codierenden Gensequenz. Ein Klon kann im Sinne der vorliegenden Erfindung auch durch eine

alleinige codierende Sequenz definiert sein., d. h. einen definierten Genotyp ohne Expressionsorganismus.

5 Diese Überführung in einen Organismus geschieht durch die bekannten Methoden der Molekularbiologie für die Transformation von Gensequenzen in Expressionsorganismen und ist abhängig von dem verwendeten Expressionsorganismus. Eine bevorzugte Methode ist die Elektroporation.

Bevorzugt erfolgt die Aufteilung in Kompartimente sofort nach der Überführung der Variantenbibliothek in den Expressionsorganismus.

Die Anzahl W_0 der Kompartiment beträgt bevorzugt zwischen 10^1 und 10^4 Kompartimenten.

Die Bibliotheksgröße B_0 dividiert durch die Kompartiment-Anzahl W_0 ergibt die Klon-Anzahl K_0 pro Kompartiment, $K_0 = B_0 / W_0$.

Jedes Kompartiment enthält eine Teilbibliothek mit einer Anzahl von K_0 Varianten der für das Biomolekül codierenden Gensequenz.

15 Besonders bevorzugt erfolgt die Aufteilung in die Kompartimente einer Mikrotiter- bzw. Deepwellplatte mit $W_0 = 96$ Kompartimenten

Nun erfolgt eine Vervielfältigung der Teilbibliotheken in den Kompartimenten durch Wachstum der Organismen oder Vervielfältigung der codierenden Sequenzen durch Matrizen-abhängige Enzyme bis zu einer Individuen-Anzahl V_0 pro Kompartiment und
20 Produktion der katalytischen Biomoleküle durch die Expressionsorganismen oder zellfreie Expressionssysteme, wie z.B. *E. coli* Lysate, Reticulocyten-Lysate, *C. lucknowese* Lysate oder Insektenzellen-Lysate.

Bevorzugt erfolgt nun eine Konservierung eines Anteils der Teilbibliothek auf
25 Organismen-Ebene oder der Ebene reiner codierender Sequenzen zum Zeitpunkt x unter Beibehaltung der Kompartiment-Zuordnung.

Die Konservierung von Organismenkulturen erfolgt bevorzugt durch die Herstellung einer 1:1 Mischung aus Organismen-Kultur und Glycerol und Lagerung dieser Mischung unter Wachstumsinhibition bei -80°C . Eine Konservierung auf Ebene der codierenden Sequenzen erfolgt durch Abnahme eines Anteils der vervielfältigten Sequenzen und Lagerung, bevorzugt bei -20°C .

Eine Bestimmung der Individuen-Anzahl $V_0(x)$ der konservierten Teilbibliothek auf Organismen-Ebene erfolgt bevorzugt durch Messung der optischen Dichte OD einer flüssigen Organismen-Kultur und Korrelation mit der Individuen-Anzahl oder Überführung eines Aliquot dieser Kultur auf ein Festmedium und Auszählen der daraus resultierenden Kolonien. Die Bestimmung der Individuen-Anzahl $V_0(x)$ der konservierten Teilbibliothek auf Ebene codierender Sequenzen erfolgt bevorzugt durch Konzentrationsbestimmung mittels spektroskopischer Methoden.

Die Individuen-Anzahl $V_0(x)$ dividiert durch die Klonanzahl pro Kompartiment K_0 ergibt den Vervielfältigungsfaktor $F_0(x)$ pro Klon, $F_0(x) = V_0(x) / K_0$.

Während oder nach dem Wachstums der Organismen oder der Vervielfältigung der Genotypen erfolgt die Produktion der Biomoleküle in den einzelnen Kompartimenten.

Im Schritt c.) des Verfahrens werden die in den einzelnen Kompartimenten enthaltenen Biomoleküle auf eine bestimmte Eigenschaft (Phänotyp), bevorzugt eine biokatalytische Aktivität, getestet.

Da jedes Kompartiment mehr als einen Klon der Variantenbibliothek enthält, lässt sich aus dem beobachteten Phänotyp keine direkten Rückschlüsse auf den Genotyp machen, da der beobachtete Phänotyp durch die Summe der im Kompartiment enthaltenen Klone resultiert.

Der für die gewünschte Eigenschaft verantwortliche Klon, der zum Beispiel die gewünschte Enzymaktivität enthält, lässt sich später aus dem Gemisch der Klone wiederfinden und isolieren.

Bevorzugt erfolgt die Test auf eine biokatalytische Aktivität durch Inkubation der in dem Kompartiment enthaltenen oder aus diesen isolierten katalytisch aktiven Biomoleküle mit entsprechenden Substraten und Zuordnung von Aktivitätswerten zu den jeweiligen Kompartimenten. Kompartimente in denen der Aktivitätswert eine definierte Schwelle überschreitet, werden als positiv gewertet.

Im Schritt d.) des Verfahrens wird mindestens ein Kompartiment ausgewählt in dem Biomoleküle enthalten sind, welche die gewünschten Eigenschaften erfüllen.

Die in diesem Kompartiments enthaltene Teilbibliothek wird nun im Schritt e.) des Verfahrens entsprechend Schritt b.) erneut in Kompartimente aufgeteilt

Vorzugsweise wird dazu die entsprechende konservierte Teilbibliothek um den Faktor $F_0(x)$ so verdünnt, dass in einem gegebenen Volumen jeder in dem Kompartiment enthaltene Klon statistisch mit der Anzahl $X = 1$ vorkommt. Dieses Volumen wird wiederum ohne vorherige Vervielfältigung auf neue Kompartimente der Anzahl W_1 aufgeteilt. Die neue Klonanzahl pro Kompartiment $K_1 = K_0 / W_1$.

15 Nun werden die Schritte c.) bis e.) des Verfahrens solange wiederholt bis die Anzahl der Klone pro Kompartiment $K_n \leq 1$. Sobald $K_n \leq 1$ erreicht ist, ist der gesuchte Phänotyp einen einzelnen Genotyp zugeordnet.

Um den Verlust von Einzelklonen und somit Varianten der Bibliothek von Biomolekülen zu verhindern, kann der Schritt e.) in der Weise ausgeführt werden, dass $X > 1$, bevorzugt: $X = 3-5$.

Der Schritt e.) kann aber auch solange wiederholt werden, bis der den gesuchten Phänotyp verursachende Klon innerhalb der neu-kompartimentierten Teilbibliothek zu finden ist. Hierbei kann X auch < 1 sein. Vorzugsweise wird bei der letzten Durchführung des Schrittes e.) die Teilbibliothek so verdünnt, dass maximal ein Klon pro Kompartiment zu finden ist und in vielen Kompartimenten kein Klon mehr enthalten ist. Damit gibt sich ein durchschnittlicher Wert von $X < 1$.

25

Anhand des nachfolgenden Ausführungsbeispiels wird die Erfindung am Beispiel der Selektion von aktiver RNase T1 aus einer Variantenbibliothek von inaktiven Varianten von RNase T1 näher beschrieben:

1. Klonierung der Gene für RNase T1 Wildtyp und His92Ala

- 5 Mit den beiden Primern A2Vo_BspHI (SEQ_ID No. 1) und A2Hi_PstI (SEQ_ID No. 2) (beide IBA Göttingen) werden die für RNase T1 Wildtyp (SEQ_ID No. 3) und die RNase T1 Variante His92Ala (SEQ_ID No. 4) codierenden Gene inklusive des Signalpeptides für eine periplasmatische Expression aus dem jeweiligen Ursprungsvektoren pA2T1 (SEQ_ID No. 5) und pA2T1_H92A (SEQ_ID No. 5, in dem SEQ_ID No. 3 durch SEQ_ID No. 4 ersetzt ist) durch eine PCR unter den nachfolgenden Bedingungen amplifiziert.

1.1 PCR:

PCR-Ansatz:	10 µl	10x VENT-Puffer (NEB, Beverly, USA)	
	2 µl	dNTPs (je 10 mM)	
	100 pmol	Primer A2Vo_BspHI	(SEQ_ID No. 1)
	100 pmol	Primer A2Hi_PstI	(SEQ_ID No. 2)
	1 µl	Ursprungsvektor (20 ng)	(SEQ_ID No. 5)
	2 U	VENT-Polymerase (NEB)	
	ad 100 µl	H ₂ O dest.	

Temperaturprofil der PCR: 2 min / 94 °C

- | | | |
|---------------|--------------------------------|--------|
| 1. | 45 sec / 94 °C (Denaturierung) | } 25 x |
| 2. | 45 sec / 57 °C (Anlagerung) | |
| 3. | 30 sec / 72 °C (Elongation) | |
| 2 min / 72 °C | | |

Die resultierenden PCR-Produkte werden mittel des QIAquick PCR-Reinigungs-Kit (Qiagen, Hilden) nach Herstellervorschrift gereinigt.

1.2 Restriktionsverdau:

Zur Klonierung der Gene in den Expressionsvektor pETBlue-2 (SEQ_ID No. 6) werden die PCR-Produkte und der Vektor mittels Restriktionsendonucleasen BspHI und PstI bzw. NcoI und PstI (alle MBI Fermentas, Vilnius, Litauen) wie folgt inkubiert:

5

Restriktionsverdau-Ansätze:

PCR-Produkte:

Vektor:

10

2 µg PCR-Produkt

4 µg pETBlue-2

2 µl 10x Puffer O⁺ (MBI)

2 µl 10x Puffer Y⁺ (MBI)

10 U BspHI

10 U NcoI

10 U PstI

10 U PstI

ad 20 µl H₂O dest.

ad 20 µl H₂O dest.

15

Die Restriktionsverdau-Ansätze werden 2 h bei 37 °C inkubiert. Zu dem „Vektor-Ansatz“ wird anschließend zur Dephosphorylierung 1 U SAP (MBI Fermentas, Vilnius, Litauen) hinzu gegeben und für weitere 30 min bei 37 °C inkubiert. Anschließend werden die Enzyme für 20 min bei 80 °C inaktiviert. Daraufhin werden die Produkte

20

mittels des QIAquick PCR-Reinigungs-Kit (Qiagen, Hilden) aufgereinigt.

1.3 Ligation, Transformation in *E. coli* und Plasmid-Reisolation

Die Vektor-DNA und das PCR-Produkt werden durch die Inkubation mit T4-DNA-Ligase wie folgt miteinander verbunden:

25

Ligase-Ansatz:

200 fmol

Vektor-DNA

600 fmol

PCR-Produkt

3 µl

10x Ligase-Puffer (MBI)

1 µl

T4-DNA-Ligase

ad 30 µl

H₂O dest.

30

Die Ansätze werden 8 h bei 16 °C inkubiert und anschließend wurde das Enzym durch 10-minütige Inkubation bei 65 °C inaktiviert. 1 µl dieses Ansatzes wurde direkt zur

Transformation kommerziell erhältlicher kompetenter ElectroTen-Zellen (Stratagene, La Jolla, USA) mittels Elektroporation eingesetzt. Die elektroporierten Zellen wurden auf Festagar-Platten mit Ampicillin ausplattiert und über Nacht bei 37 °C kultiviert. Ausgehend von einer resultierenden Einzelkolonie wird das fertige Plasmid mittels des
5 Plasmid-Reinigungskits QIAprep Minipräparations-Kit (Qiagen, Hilden) nach Herstellervorschrift reisoliert.

1.4 Herstellung einer Plasmid-Mischung als RNase T1-Testbibliothek:

Als Resultat aus den vorangegangenen Schritten werden die beiden Plasmide pETBlue-RNaseT1-Wildtyp und pETBlue-RNaseT1-His92Ala erhalten.

- 10 Zur Herstellung einer Testbibliothek werden die Plasmide wie folgt gemischt:
1 pg pETBlue-RNaseT1-Wildtyp wird mit 1 µg pETBlue-RNaseT1-His92Ala gemischt. Dadurch erhält man ein Verhältnis von 1 : 1.000.000 aus RNase T1 Wildtyp (aktiv) und der Variante His92Ala (inaktiv).

1.5 Herstellung des Expressionsstammes:

- 15 Für die Expression der RNase T1-Testbibliothek wird ein *E. coli* Stamm benötigt, bei dem die RNase I ausgeschaltet ist. Entsprechende Stämme, wie z. B. AT9 (*rna-19* λ *gdhA2 relA1 spoT1 metB1*) sind über das *E. coli* Genetic Stock Center New Haven, USA verfügbar. Der im Beispiel verwendete Expressionsvektor pETBlue-2 benötigt zusätzlich die T7-RNA-Polymerase für die Expression, welche in *E. coli* nicht
20 vorhanden ist. Mit dem kommerziell erhältlichen λ DE3-Lysogenisierungs-Kit (Novagen, Madison, USA) wird nach Herstellervorschrift das T7-RNA-Polymerase codierenden Gen in den *E. coli* Stamm AT9 eingeführt. Hierdurch erhält man einen *E. coli* Stamm, der sich durch die Abwesenheit von RNase I und das Vorhandensein der T7-RNA-Polymerase (DE3) auszeichnet. Von diesem Stamm wurden mittels Standard-
25 Molekularbiologie-Methoden elektrokompente Zellen hergestellt und bei -80 °C gelagert.

1.6 Transformation des Expressionsstammes mit der Testbibliothek:

- In den wie oben beschrieben hergestellten Expressionsstamm wird 1 ng der Plasmid-Mischung als Testbibliothek mittels Elektroporation transformiert und die
30 resultierenden Zellen nach 1-stündigem Wachstum bei 37 °C in 10 ml Flüssigmedium

(LB-Medium: 10 g Trypton, 5 g Hefe-Extrakt (alles Becton Dickinson, Heidelberg), 10 g NaCl (Sigma, Deisenhofen)) mit Ampicillin aufgenommen

Die so erhaltene Vorkultur wird sofort auf eine 96-Well Mikrotiterplatte (MTP) aufgeteilt (je 100 µl pro well) und über Nacht bei 30 °C und 800 Upm inkubiert.

- 5 Durch die Transformationen mittels Elektroporation werden ca. 3 Millionen transformierte Klone erhalten.

1.7 Wachstum der Hauptkultur und Expression von RNase T1

- 10 Eine 96er Deepwellplatte (DWP) wird mit jeweils 1,5 ml Flüssigmedium mit Ampicillin pro Well befüllt. Das Medium wird mit jeweils 50 µl aus der Vorkultur-MTP beimpft und die DWP bei 37 °C und 800 Upm kultiviert. Beim Erreichen einer optischen Dichte OD_{600} der Kulturen von $OD_{600} = 1,0$ werden die Kulturen mit 1 mmol/Liter IPTG induziert. Anschließend wird die Platte für weitere 4 h bei 37 °C und 800 Upm inkubiert.

1.8 Präparation der Protein-Proben

- 15 Durch das Signalpeptid ompA werden die exprimierten RNase T1 – Moleküle in den periplasmatischen Raum des Expressionsbakteriums geleitet. Durch einen osmotischen Schock können die Proteine sehr leicht präpariert werden. Die Reinigungsprozedur umfasst dabei folgende Schritte:

- Sammeln der Zellen durch Zentrifugation bei 4000 Upm, 4 °C für 5 min
- 20 • Abschütten des Medien-Überstandes
- Resuspension des Bakterienpellets in jeweils 25 µl Puffer A (50 mmol/Liter Tris/HCl, pH 7,5, 10 mmol/Liter EDTA, 15 % Saccharose w/v)
- Inkubation auf Eis für 30 min
- Zugabe von jeweils 125 µl Puffer B (50 mmol/Liter Tris/HCl, pH 7,5, 10 mmol/Liter EDTA)
- 25 • Zentrifugation bei 4000 Upm, 4 °C, für 20 min
- Abnahme des Überstandes und Überführung in eine MTP (Periplasma)
- Aufbewahren der Bakterienpellets

1.9 Herstellung des Substrates für RNase T1

Als Substrat dient ein doppelsträngiges DNA-Molekül mit zentralem einzelsträngigen Bereich, welcher einen Guanosin-RNA-Baustein als Angriffspunkt für das Enzym enthält. Die Enden dieses Substrates sind mit unterschiedlichen Farbstoffen für den roten (Cy5) und den grünen (RhG) Spektralbereich markiert. Um ein Ausbleichen der markierten Substrate zu verhindern, werden die entsprechenden Lösungen und Inkubationsansätze stets vor Licht geschützt. Die Puffer und Ansätze werden mit DEPC-behandeltem Wasser hergestellt. Das Substrat setzt sich aus den folgenden 3 Oligonucleotiden (IBA Göttingen) zusammen:

1. Sub_G:

5'-Cy5-CCATACCAGCCAGCCACAArGCAAGCCACCGAAGCACAGATA-RhG-3'

2. T1_Sub_Li:

5'-GTGGCTGGCTGGTATGGA-3' (SEQ_ID No. 7)

3. T1_Sub_Re:

5'-TATCTGTGCTTCGGTGGC-3' (SEQ_ID No. 8)

Durch die nachfolgend beschriebene Hybridisierung werden die 3 Bestandteile wie folgt zu einem doppelsträngigen Substrat aneinander angelagert:

Hybridisierungs-Ansatz:

1000 pmol Sub_G
1200 pmol T1_Sub_Li
1200 pmol T1_Sub_Re
20 µl MES (1 mol/Liter, pH 6,0)
ad 1000 µl DEPC-H₂O

Hybridisierungs-Programm:

1. 10 s, 94°C;
2. Abkühlen auf 25 °C mit 0,1 °C/s
3. 4 °C

1.10 Inkubation der Protein-Proben mit dem Substrat

In einer MTP werden jeweils 10 µl des doppelsträngigen Substrates pro Well vorgelegt. Dazu werden jeweils 10 µl der aus dem Periplasma isolierten Protein-Proben hinzugegeben, die MTP luftdicht verschlossen und dunkel für 24 h bei 37 °C inkubiert. Anschließend werden jeweils 5 µl dieser Ansätze in eine MTP mit Glasboden überführt und mit jeweils 250 µl Puffer C (100 mmol/Liter MES, pH 6,0, 100 mmol/Liter NaCl, 2 mmol/Liter EDTA) gemischt.

1.11 Aktivitätsbestimmung

Zur Bestimmung der Enzymaktivität wird die Platte mit Glasboden, in die die Inkubationsansätze wie unter 1.10 beschrieben überführt wurden, am Fluoreszenz-Korrelations-Spektroskop ConfoCor 2 (Evotec Biosystems, Hamburg und Carl Zeiss Microscopy, Jena) vermessen. Die Auswertung der Daten erfolgt mit der ConfoCor 2-Software (Version 2.5).

Für die Messungen wird ein Argon-Laser ($\lambda = 488 \text{ nm}$) zur Anregung von RhG in Kombination mit einem Helium/Neon-Laser ($\lambda = 633 \text{ nm}$) für Cy5 eingesetzt. Das FCS-Messvolumen in den Kavitäten wurde $200 \text{ }\mu\text{m}$ über der Glasbodenoberfläche justiert.

Die Messungen erfolgen für 20 s pro Well.

Durch eine Kreuzkorrelationsanalyse der erhaltenen Daten kann auf eine eventuelle Spaltung des Substrates geschlossen werden. Eine Spaltung des Substrates durch RNase T1 führt zu einer Entkopplung der beiden Fluoreszenzfarbstoffe und somit zum Verlust des Kreuzkorrelationssignales. Ungeschnittene Substrat-Moleküle tragen hingegen beide Farbstoffe und liefern ein hohes Signal.

Durch die Aufteilung der durch die Transformation erhaltenen 3 Millionen Klone und das Mischungsverhältnis zwischen aktiver RNase T1 Wildtyp und inaktiver RNase T1 His92Ala von 1 : 1.000.000 sollten theoretisch 3 Wells mit Aktivität durch die Messungen detektierbar sein. Statistische Abweichungen zwischen 1-5 Wells mit Aktivität sind jedoch möglich.

Fig. 1 zeigt die so erhaltenen Messdaten für ein nach Ausführungsbeispiel 1 hergestellte RNase T1-Testbibliothek bestehend aus 3 Millionen Klonen auf einer Platte mit einem Mischungsverhältnis von RNase T1 Wildtyp und RNase T1-His92Ala von 1 : 1.000.000. Die RNase T1-Aktivität wurde wie oben beschrieben mittels Kreuzkorrelationsanalyse detektiert. Für eine bessere Übersicht wurde eine reziproke Darstellung gewählt, d.h. hohe Peaks bedeuten ein niedriges Signal und niedrige Peaks ein hohes Signal. Fig. 1 zeigt 2 deutliche Peaks, die durch einen Verlust des Kreuzkorrelationssignales hervorgerufen werden. Diese beiden Peaks zeigen, dass in dem Experiment eine RNase T1-Aktivität in zwei von 96 Wells sicher vorhanden war.

2. Reisolation der Teilbibliothek

In der nach Ausführungsbeispiel 1 erhaltenen Platte mit den aufbewahrten Bakterienpellets aus der Proteinpräparation wird in einem der Wells, welche in der Aktivitätsbestimmung (1.11) eine RNase T1-Aktivität gezeigt hat, eine Plasmidpräparation mittels dem QIAprep Minipräparations-Kit (Qiagen, Hilden) durchgeführt.

Durch die ursprüngliche Aufteilung von 3 Millionen Klonen auf die Platte ergab sich eine Anzahl von $3.000.000 / 96 = 31.250$ unterschiedlichen Klonen pro Well. Es besteht somit in der isolierten Teilbibliothek ein Mischungsverhältnis von RNase T1 Wildtyp zu RNase T1 His92Ala von 1 : 32.250.

2.1 Weitere Vereinzelungen

Durch eine Transformation von verschiedenen Aliquots der so erhaltenen Teilbibliothek analog der vorgehensweise von Ausführungsbeispiel 1.6 wurde die Menge von Plasmid-DNA bestimmt, welche notwendig ist, um nun ca. 100.000 transformierte Klone mittels Elektroporation zu erhalten.

Anschließend wurde die bestimmte Menge der Teilbibliothek in den Expressionsstamm transformiert und das gleiche Verfahren wie für die ursprüngliche Testbibliothek durchlaufen. Da ca. 100.000 Klone aufgeteilt wurden und das neue Mischungsverhältnis 1 : 32.250 betrug, waren wieder theoretisch 3 Wells mit detektierbarer Aktivität zu erwarten.

Die Plasmide wurden wiederum in einer der Wells mit Aktivität aus dem Bakterienpellet reisoliert. Das Mischungsverhältnis in dieser erneut angereicherten Teilbibliothek war nun $100.000 / 96 = 1050$.

Eine weitere Wiederholung des dargestellten Schemas mit einer Aufteilung von jetzt ca. 3000 Klonen ergab eine nochmals angereicherte Teilbibliothek mit einem Mischungsverhältnis von $3000 / 96 = 31$.

Indem von dieser letzten Teilbibliothek 96 Klone auf eine MTP aufgeteilt wurden, resultierten daraus etwa 3 Wells mit Aktivität. Da diese Aktivitäten nun jeweils von einem vereinzelteten Klon verursacht wurden, konnte diesem die Aktivität von RNase T1 Wildtyp direkt zugeordnet werden.

Abkürzungsverzeichnis:

In der Erfindungsbeschreibung werden folgende Abkürzungen verwendet:

	<i>B. subtilis</i>	<i>Bacillus subtilis</i>
5	<i>C. lucknowese</i>	<i>Chrysosporium lucknowese</i>
	Cy5	Fluoreszenzfarbstoff Cy5™ (Amersham Biosciences UK Limited, Little Chalfont, Buckinghamshire, GB)
	DEPC	Diethylpyrocarbonat
	DWP	Deepwellplatte
10	<i>E. coli</i>	<i>Escherichia coli</i>
	EDTA	Ethylendiamintetraessigsäure
	h	Stunde
	IPTG	Isopropyl-β-D-thiogalacto-pyranosid
	LB	Luria Broth
15	MES	Morpholinoethansulfonsäure
	min	Minuten
	MTP	Microtiterplatte
	OD	optische Dichte
	OD ₆₀₀	optische Dichte bei 600 nm
20	ompA	äußeres Membranprotein A aus <i>E. coli</i>
	p	Plasmid
	PCR	Polymerase-Kettenreaktion
	PT7	T7-Promotor
	rG	Guanylsäurerest
25	Upm	Umdrehungen pro Minute
	RhG	Rhodamin Grün (Fluoreszenzfarbstoff)
	SAP	Alkalische Phosphatase aus Krabben
	<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i> (Hefe)
	Tris	Tris-(hydroxymethyl)-aminomethan
30	T4	vom Bakteriophage T4 abstammend
	U	Unit (Einheit für Enzymaktivität)
	w/v	bei Prozentangaben: Gewicht (w = weight) pro Volumen (v)

Patentansprüche

1. Verfahren zur Selektion von Biomolekülen aus Varianten-Bibliotheken von Biomolekülen umfassend die Schritte:

a.) Herstellung einer Varianten-Bibliothek, bestehend aus einer Anzahl (B_0) von Varianten, der für das Biomolekül codierenden Gensequenz, und

b.) Aufteilung der Variantenbibliothek in eine Anzahl von Kompartimenten (W_0), die mindestens um einen Faktor 10, bevorzugt einen Faktor 100, kleiner ist, als die Anzahl (B_0) der in der Varianten-Bibliothek enthaltenen Varianten,

wobei jedes Kompartiment eine Teilbibliothek enthält, die $K_0 = B_0 / W_0$ Varianten enthält,

c.) Produktion von Biomolekülen in den Kompartimenten, und Test der in den einzelnen Kompartimenten erhaltenen Biomoleküle auf eine bestimmte Eigenschaft, bevorzugt eine biokatalytische Aktivität,

d.) Auswahl mindestens eines Kompartiments, in dem Biomoleküle enthalten sind, welche die gewünschten Eigenschaften, bevorzugt eine biokatalytische Aktivität, erfüllen,

e.) Aufteilung der in dem ausgewählten Kompartiment enthaltenen Teilbibliothek in weitere Kompartimente und n-faches Wiederholen der Schritte c.) bis e.) bis in jedem Kompartiment nur noch maximal eine Variante ($K_n \leq 1$) der für das Biomolekül codierenden Gensequenz enthalten ist.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Variantenbibliothek 10^3 bis 10^{15} Varianten, bevorzugt 10^5 bis 10^9 Varianten, der Gensequenz des Biomoleküls enthält.

3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet**, dass die Variantenbibliothek im Schritt b.) auf 10^1 bis 10^4 , bevorzugt 96, Kompartimente aufgeteilt wird.
- 5 4. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, dass die Variantenbibliothek vor der Aufteilung im Schritt b) in einen Organismus überführt wird.
- 10 5. Verfahren nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet**, dass die Kultur des Organismus nach der Aufteilung im Schritt c.) auf eine Organismenzahl von 10^8 bis 10^9 pro Kompartiment vervielfältigt wird.
- 15 6. Verfahren nach Anspruch 4 oder 5, **dadurch gekennzeichnet**, dass die Organismen auch die Produktion der Biomoleküle durchführen.
- 20 7. Verfahren nach einem der Ansprüche 4 bis 6, **dadurch gekennzeichnet**, dass die Teilbibliotheken in den Kompartimenten aus den Organismen reisoliert werden und die Produktion der Biomoleküle durch zellfreie Systeme, durchgeführt wird.
- 25 8. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, dass die Vervielfältigung der Teilbibliotheken und die Produktion der Biomoleküle durch zellfreie Systeme durchgeführt wird.
- 30 9. Verfahren nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, dass die Variantenbibliothek aus DNA-Plasmiden besteht, welche die für das Biomolekül codierende Gensequenz enthalten.
10. Verfahren nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet**, dass die Variantenbibliothek aus linearen Nukleinsäuremolekülen besteht, welche die für das Biomolekül codierende Gensequenz enthalten.

11: Verfahren nach einem der Ansprüche 1 bis 10, **dadurch gekennzeichnet, dass** die Biomoleküle Enzyme oder Ribozyme oder andere Biomoleküle sind, welche eine biokatalytische Aktivität besitzen.

5 12: Verfahren nach einem der Ansprüche 1 bis 11, **dadurch gekennzeichnet, dass** der Test auf eine biokatalytische Aktivität mittels physikalischer Messmethoden, wie vorzugsweise der UV/VIS-Spektroskopie, der Fluoreszenz-Spektroskopie oder der Fluoreszenz-Korrelations-Spektroskopie, erfolgt.

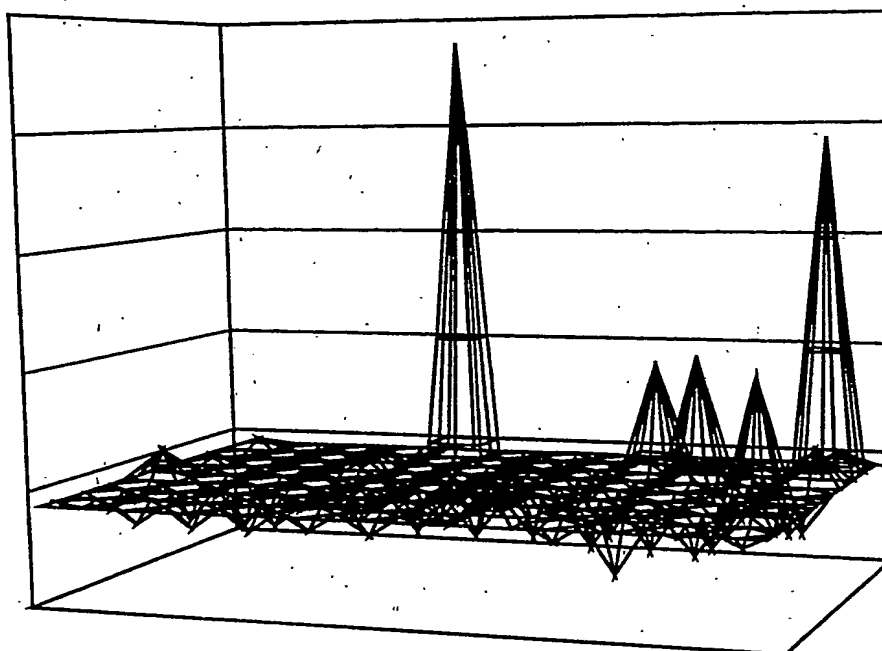


Fig. 1

SEQUENZPROTOKOLL - SEQUENCE LISTING

<110> Universität Leipzig
 <120> Verfahren zur Selektion von Biomolekülen aus Varianten-
 Bibliotheken von Biomolekülen
 <130> 401P03DE
 <160> 8
 <170> PatentIn version 3.1

<210> 1
 <211> 28
 <212> DNA
 <213> artificial
 <400> 1
 caattctgca gttgcgttca cgtcgttg

28

<210> 2
 <211> 28
 <212> DNA
 <213> artificial
 <400> 2
 taaggctcat gaaaaacaca gctatcgc

28

<210> 3
 <211> 378
 <212> DNA
 <213> Escherichia coli
 <220>
 <221> ompA-Signalpeptid
 <222> (1)..(63)
 <223>
 <220>
 <221> RNase T1 Wildtyp
 <222> (64)..(378)
 <223>

<400> 3
 atgaaaaaca cagctatcgc gattgcagtg gcactggctg gtttcgctac cgtagcgcag 60
 gccgcatgcg actacacttg cggttctaac tgctactctt cttcagacgt ttctactgct 120
 caggcggccg gatataaact tcacgaagac ggtgaaactg ttggatccaa ttcttaccba 180
 cacaagtaca acaactacga aggttttgat ttctctgtga gctctcccta ctacgaatgg 240
 cctatcctct cgagcgggtga tgtttactct ggtgggtccc cgggtgctga ccgtgtcgtc 300
 ttcaacgaaa acaaccaact agctgggtgtt atcactcaca ctgggtgcttc tggttaacaac 360
 ttcgttgaat gtacataa 378

<210> 4
 <211> 378
 <212> DNA
 <213> Escherichia coli
 <220>
 <221> ompA-Signalpeptid
 <222> (1)..(63)

```

<223>
<220>
<221> RNaseT1-His92Ala
<222> (64)..(378)
<223>
<400> 4
atgaaaaaca cagctatcgc gattgcagtg gcactggctg gtttcgctac cgtagcgcag      60
gccgcatgcg actacacttg cggttctaac tgctactctt cttcagacgt ttctactgct      120
caggcggccg gatataaact tcacgaagac ggtgaaactg ttggatccaa ttcttaacca      180
cacaagtaca acaactaaga aggttttgat ttctctgtga gctctcccta ctacgaatgg      240
cctatcctct cgagcgggta tgtttactct ggtgggtccc cgggtgctga ccgtgtcgtc      300
ttcaacgaaa acaaccaact agctgggtgt atcactgcc a ctggtgcttc tggttaacaac      360
ttcgttgaat gtacataa      378

```

```

<210> 5
<211> 7336
<212> DNA
<213> Plasmid pA2T1
<220>
<221> lac Promotor
<222> (1)..(371)<223>
<220>
<221> ompA-Signalpeptid
<222> (393)..(455)
<223>
<220>
<221> RNaseT1-Wildtyp
<222> (456)..(770)
<223>
<220>
<221> lacI-Gen
<222> (1664)..(2887)
<223>
<220>
<221> ORI
<222> (4924)..(5115)
<223>
<220>
<221> Beta-Lactamase (Amp)
<222> (7165)..(6302)
<223>

```

```

<400> 5
taggcgtatc acgaggccct ttggataacc agaagcaata aaaaatcaaa tcggatttca      60
ctatataatc tcaactttatc taagatgaat ccgatggaag catcctgttt tctctcaatt      120
tttttatcta aaaccagcg ttcgatgctt ctttgagcga acgatcaaaa ataagtgcct      180
tccatcaaaa aaaatattct caacataaaa aactttgtgt aatacttgta acgctacatg      240
gagattaact caatctagct agagaggctt tacactttat gcttccggct cgtataatgt      300
gtggaattgt gagcggataa caatttcaca caggaaacag ctatgaccat gattacggat      360

```

tcactggaac tctagataac gaggcgcaaa aaatgaaaaa cacagctatc gcgattgcag 420
 tggcactggc tggtttccgt accgtagcgc aggcgcgcgt cgactacact tgtggttcca 480
 actgctactc ttcttcagac gtttctactg ctcaagcggc cggatataaa cttcacgaag 540
 acggtgaaac tgttgatcc aattcttacc cacacaaata caacaactac gaagggtttg 600
 atttctctgt gagctctccc tactacgaat ggcctatcct ctcgagcggg gatgtttact 660
 ctggtgggtc cccgggtgct gaccgtgtcg tcttcaacga aaacaaccaa ctagctgggtg 720
 ttafcaactca cactggtgct tctggttaaca acttcgttga atgtacataa gcttggtatcg 780
 atccgggctg agcaacgacg tgaacgcaat gcgttccgac gttcaggctg ctaaagatga 840
 cgcagctcgt gctaaccagc gtctggacaa catggctact aaataccgca agtaatagta 900
 cctgtgaagt gaaaaatggc gcacattgtg cgacattttt tttgtctgcc gtttaccgct 960
 actgcgtcac gcgtaacata ttcccttgct ctggttcacc attctgcgct gactctactg 1020
 aaggcgcatt gctggctgcg ggagttgctc cactgctcac cgaaaccgga taccctgccc 1080
 gacgatacaa cgctttatcg actaacttct gatctacagc cttattgtct ttaaattgcg 1140
 taaagcctgc tggcagtgtg tatggcattg tctgaacggt ctgctgttct cctgccgata 1200
 gtggtcgtatg tacttcaaca taacgcattc cgttaggctc cacggaatat ttcaccggtt 1260
 cgttgatcac tttcaccggc gttcccgctc gcacgctgga gaacaaggct ttaatatccg 1320
 gtgcattcat gcgaatacac cctgaactga cgcgcaaacc gacgctgtcc ggcgactgg 1380
 taccatgaat gaggtattcg ccattaccat gcgcgaggcg cagtgcgtaa cgtcctagcg 1440
 ggttatttgg tccggcagga acgactggcg gtaatttaat gccacgctcc agcgaacgct 1500
 gacgaatgcc tgcgtaggc gtccagggtg ggtagggat tttctgcca acacgcgttt 1560
 ccatcaccgg cgtttccagc ccctgcaatc caatacctat tggataaacc tgcacaatat 1620
 tttctcccg cgataataa taaaggcgca gctctgcaag gttgacacca tcgaatggcg 1680
 caaaaccttt cgcggtatgg catgatagcg cccggaagag agtcaattca ggggtggtgaa 1740
 tgtgaaacca gtaacggttat acgatgtcgc agagtatgcc ggtgtctctt atcagaccgt 1800
 ttcccgctg gtgaaccagg ccagccacgt ttctgcgaaa acgcgggaaa aagtggaagc 1860
 ggcgatggcg gagctgaatt acattcccaa ccgctgggca caacaactgg cgggcaaaca 1920
 gtogttgctg attggcggtt ccacctccag tctggccctg cagcgccgt cgcaaattgt 1980
 cgcggcgatt aaatctcgcg ccgatcaact ggggtgccagc gtggtggtgt cgatggtaga 2040
 acgaagcggc gtcgaagcct gtaaagcggc ggtgcacaat cttctcgcgc aacgcgtcag 2100
 tgggctgatc attaaactat cgctggatga ccaggatgcc attgctgtgg aagctgcctg 2160
 cactaatgtt ccggcggtat ttcttgatgt ctctgaccag acacccatca acagtattat 2220

tttctcccat gaagacggta cgcgactggg cgtggagcat ctggtcgcat tgggtcacca 2280
 gcaaatcgcg ctgttagcgg gccattaag ttctgtctcg gcgcgtctgc gtctggctgg 2340
 ctggcataaa tatctcactc gcaatcaaat tcagccgata gcggaacggg aaggcgactg 2400
 gagtgccatg tccggttttc aacaaacat gcaaatgctg aatgagggca tcgttcccac 2460
 tgcgatgctg gttgccaacg atcagatggc gctgggcgca atgcgcgcca ttaccgagtc 2520
 cgggctgcgc gttggtgcgg atatctcggg agtgggatac gacgataccg aagacagctc 2580
 atgttatata ccgcgcgcaa ccaccatcaa acaggatttt cgctgtctgg ggcaaaccag 2640
 cgtggaccgc ttgctgcaac tctctcaggg ccaggcgggtg aagggaatc agctgttgcc 2700
 cgtctcactg gtgaaaagaa aaaccaccct ggcgcccaat acgcaaaccg cctctccccg 2760
 cgcgttgccc gattcattaa tgcagctggc acgacagggt tcccgactgg aaagcgggca 2820
 gtgagcgcaa cgcaattaat gtgagttagc tcactcatta ggcaccccag gctttacact 2880
 ttatgctaac gataatcccc tgacgcgggtg catcaggtaa taacagttgt gaaggaatag 2940
 ttatcgctgt accaggtttt ggcaccgggg cgatagtgtt attggcttca aggatcaaca 3000
 ttgccgcagt atcaaaacgt cgggcaatag cctgaagggt tttatcccct tcttgaccg 3060
 tatacgtttg atttgccc accagtcggc ttccggttg tggtagcgga taatcaaccg 3120
 cccaggcagc ctggatggcg ctaaaagcgc cgataagcgt gagtgtaagc aaagacgcgc 3180
 gtttcattgt aaacctcctg tatttgccgg agactcacgc tgaaacgtcg gatggcgctt 3240
 atgttcacct gaaacaaaa cactcctgtg caggtcagtg taaacattga ccatccggca 3300
 atgtgagcca accggatgaa agctgtcctt ttagtttagc taagtgcagc ggctttggcg 3360
 cgaattgcgc gaatcatcgc ttccagacct tgtgaacgag atggggtgag atgttggggtg 3420
 agcgccattt tttcaaacca cggacgcaca tcgaaattga caatatcctg cggcgtcac 3480
 tgatcgtaga gaataaagac gaccgcaata agccctttca caatcgccgc atcgctgtcg 3540
 ccctgtaatt caataattcc ctgggcattc tggcgcatga caatccacac ctgactctga 3600
 cagccctgaa tgcatttttg tggacttctg tcttcgtcgc gtaattctgg cagacgctgg 3660
 gggaccgatg cccttgagag ccttcaacc agtcagctcc ttccggtggg cgcggggcat 3720
 gactatcgtc gcgcactta tgactgtctt ctttatcatg caactcgtag gacaggtgcc 3780
 ggcagegctc tgggtcattt tcggcgagga ccgctttcgc tggagcgcgga cgatgatcg 3840
 cctgtcgctt gcggtattcg gaattctgca cgccctcgct caagccttcg tcaactggctc 3900
 cgccacaaa cgtttcggcg agaagcaggc cattatcgcc ggcattggcg cgcacgcgct 3960
 gggctacgtc ttgctggcgt tcgcgacgcg aggctggatg gccttcccca ttatgattct 4020

tctcgtttcc	ggcggcatcg	ggatgcccgc	gttgaggcc	atgctgtcca	ggcaggtaga	4080
tgacgaccat	cagggacagc	ttcaaggatc	gctcgggct	cttaccagcc	taacttcgat	4140
cactggaccg	ctgatcgtca	cggcgattta	tgccgcctcg	gcgagcacat	ggaacgggtt	4200
ggcatggatt	gtaggcgccg	ccctatacct	tgtctgcctc	cccgcgttgc	gtcgcgggtgc	4260
atggagccgg	gccacctcga	cctgaatgga	agccggcggc	acctcgctaa	cggattcacc	4320
actccaagaa	ttggagccaa	tcaattcttg	cggagaactg	tgaatgcgca	aaccaaccct	4380
tggcagaaca	tatccatcgc	gtccgccatc	tccagcagcc	gcacgcggcg	catctcgggc	4440
agcgttgggt	cctggccacg	ggtgcgcgatg	atcgtgctcc	tgtcgttgag	gacccggcta	4500
ggctggcggg	gttgcccttac	tggttagcag	aatgaatcac	cgatacgca	gcgaacgtga	4560
agcgactgct	gctgcaaaac	gtctgcgacc	tgagcaacaa	catgaatggt	cttcggtttc	4620
cgtgtttcgt	aaagtctgga	aacgcggaag	tcagcgccct	gcaccattat	gttcgggatc	4680
tgcatcgcag	gatgctgctg	gctaccctgt	ggaacaccta	catctgtatt	aacgaagcgc	4740
tggcattgac	cctgagtgat	ttttctctgg	tccgcgcgca	tccataccgc	cagttgttta	4800
ccctcacaac	gttcacagta	ccgggcatgt	tcatcatcag	taaccggtat	cgtgagcatc	4860
ctctctcgtt	tcatcggtat	cattaccccc	atgaacagaa	atccccctta	cacggaggca	4920
tcagtgacca	aacaggaaaa	aaccgccctt	aacatggccc	gctttatcag	aagccagaca	4980
ttaacgcttc	tggagaaact	caacgagctg	gacgcggatg	aacaggcaga	catctgtgaa	5040
tcgcttcacg	accacgctga	tgagctttac	cgcagctgcc	tcgcgcgttt	cggtgatgac	5100
ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	cagcttgtct	gtaagcggat	5160
gccgggagca	gacaagcccc	tcagggcgcg	tcagcgggtg	ttggcgggtg	tcggggcgca	5220
gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	gcggcatcag	5280
agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	tgcgtaagga	5340
gaaaataacc	catcaggcgc	tcttcgcgtt	cctcgtctac	tgactcgtcg	cgctcggctg	5400
ttcggctgcg	gcgagcggtg	tcagctcact	caaaggcggg	aatacgggta	tccacagaat	5460
caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	5520
aaaaggccgc	gttgctggcg	ttttccata	ggctccgccc	ccctgacgag	catcacaaaa	5580
atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	5640
cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	5700
ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	5760
gttcggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccc	5820
accgctgcgc	cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	5880

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcgggtgcta 5940
cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct 6000
gcgctctgct gaagccagtt accttoggaa aaagagttgg tagctcttga tccggcaaac 6060
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 6120
aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaacgaaa 6180
actcacgtta agggattttg gtcacgagat tatcaaaaag gatcttcacc tagatccttt 6240
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggctctgaca 6300
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatct cgttcatcca 6360
tagttgcttg actccccgtc gtgtagataa ctacgatacg ggagggttta ccatctggcc 6420
ccagtgtctg aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa 6480
accagccagc cggaagggcc gagcgagaa gtggctctgc aactttatcc gcctccatcc 6540
agtctattaa ttgttgccgg gaagetagag taagtagttc gccagttaat agtttgcgca 6600
acgttggtgc cattgctgca ggcacgtgg tgtcacgctc gtcgtttggt atggcttcat 6660
tcagctccgg ttcccaacga tcaaggcgag ttacatgac ccccatgttg tgcaaaaaag 6720
cggtagctc ctctggctct ccgacgttg tcagaagtaa gttggccgca gtgttatcac 6780
tcattggtat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 6840
ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt 6900
gctettgccc ggcgtcaaca cgggataata ccgcgccaca tagcagaact taaaagtgc 6960
tcattcattg aaaaagttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 7020
ccagttcgat gtaaccact cgtgcaccca actgatcttc agcatctttt actttcacca 7080
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca 7140
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 7200
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 7260
ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt attatcatga 7320
cattaaccta taaaaa 7336

<210> 6
<211> 3653
<212> DNA
<213> Plasmid pETBlue-2
<220>
<221> T7-Promotor
<222> (1)..(17)
<223>

```

<220>
<221> lac Operator
<222> (22)..(42)
<223>
<220>
<221> f1 ORI
<222> (1096)..(1551)
<223>
<220>
<221> Beta-Lactamase (Amp)
<222> (2556)..(1669)
<223>
<220>
<221> pUC ORI
<222> (3206)..(3250)
<223>
<220>
<221> lac Operator
<222> (3606)..(3625)
<223>
<400> 6
taatacgact cactataggg gaattgtgag cggataacaa ttcccctcta gacttacaat 60
ttccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtagc ggcctcttcg 120
ctattacgcc agcttgcgaa cggtaggtgc gctgcaaggc gattaagttg ggtaacgcca 180
ggattctccc agtcacgacg ttgtaaaacg acggccagcg agagatcttg attggctagc 240
agaataattt tgtttaactt taagaaggag atataccatg gcgatatccc gggagctcgt 300
ggatccgaat tctgtacagg cgcgcctgca ggacgtcgac ggtaccatcg atacgcgttc 360
gaagcttgcg gccgcacagc tgtatacacg tgcaagccag ccagaactcg ctctgaaga 420
cccagaggat ctcgagcacc accaccacca ccactaatgt taattaagtt gggcgttgta 480
atcatagtca taatcaatac tcctgactgc gttagcaatt taactgtgat aaactaccgc 540
attaaagcta ttcgatgata agctgtcaaa catgataatt cttgaagacg aaagggccta 600
ggctgataaa acagaatttg cctggcgcca gtagcgcggt ggtccacct gaccccatgc 660
cgaactcaga agtgaaacgc cgtagcgccg atggtagtgt ggggtctccc catgcgagag 720
tagggaactg ccaggcatca aataaaacga aaggctcagt cgaaagactg ggcctttcgt 780
tttatctgtt gtttgtcggg gaacgctctc ctgagtagga caaatccgcc gggagcggat 840
ttgaacggtg cgaagcaacg gcccggaggg tggcgggcag gacgcccgcc ataaactgcc 900
aggcatcaaa ttaagcagaa ggccatcctg acggatggcc tttttgcgtt tctacaaact 960
cttttgttta tttttctaaa tacattcaaa tatgtatccg ctgagcaata actagcataa 1020
ccccttgggg cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatatcc 1080
ggattggcga atgggacgcg ccctgtagcg gcgcattaag cgcggcggtt gtggtggtta 1140
cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc 1200

```

cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctccctt 1260
tagggttccg atttagtgct ttacggcaccc tcgaccccaa aaaacttgat tagggtgatg 1320
gttcacgtag tgggccatcg cctgataga cggtttttcg ccctttgacg ttggagtcca 1380
cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcggctc 1440
attcttttga tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga 1500
tttaacaaaa atttaacggg aattttaaca aaatattaac gtttacaatt tctggcggca 1560
cgatggcatg agattatcaa aaaggatcct cacctagatc cttttaaatt aaaaatgaag 1620
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat 1680
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc 1740
cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat 1800
accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag 1860
ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg 1920
ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc 1980
tacaggcatc gtggtgtcac gctcgtcgtt tggatatggct tcattcagct ccggttccca 2040
acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcgggta gtccttcgg 2100
tcctccgacg gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc 2160
actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta 2220
ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc 2280
aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg 2340
ttcttcgggg cgaaaactct caaggatcct accgctgttg agatccagtt cgatgtaacc 2400
cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc 2460
aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatggtgaat 2520
actcatactc ttcttttttc aatcatgacc aaaatccctt aacgtgagtt ttogttccac 2580
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc 2640
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cgggtggttg tttgccggat 2700
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat 2760
actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct 2820
acatacctcg ctctgctaata cctgttacca gtggctgctg ccagtggcga taagtcgtgt 2880
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg 2940
gggggttcgt gcaçacagcc cagcttggag cgaacgacct acaccgaact gagataccta 3000
cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg 3060

gtaagcggca gggtcggaac aggagagcgc acgaggggagc ttccaggggg aaacgcctgg 3120
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc 3180
tcgtcagggg ggcgggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg 3240
gccttttgct ggcccttttg tcacatgttc tttcctgcgt tatccctga ttctgtggat 3300
aacgtatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc 3360
agcgagtcag tgagcgagga agccggcgat aatggcctgc ttctcgccga aacgtttggt 3420
ggcgggacca gtgacgaagg cttgagcgag ggcgtgcaag attccgaata ccgcaagcga 3480
caggccgatc atcgtcgcgc tccagcgaaa gcggtcctcg ccgaaaatga cccagagcgc 3540
tgccggcacc tgtcctacga gttgcatgat aaagaagaca gtcataagtg cggcgacgac 3600
cgggtgaattg tgagcgctca caattctcgt gacatcataa cgtcccgcga aat 3653

<210> 7
<211> 18
<212> DNA
<213> artificial
<400> 7
gtggctggct ggtatgga

18

<210> 8
<211> 18
<212> DNA
<213> artificial
<400> 8
tatctgtgct tcggtggc

18